On hierarchical competition in oligopoly

Ludovic A. Julien, Olivier Musy, Aurélien Saidi

On hierarchical competition in oligopoly

Ludovic A. Julien, Olivier Musy, Aurélien W. Saïdi

1 Introduction

Stackelberg competition has been widely studied in the literature. Several extensions have been added to the original duopoly model (Stackelberg 1934), including a larger number of stages (Boyer and Moreaux 1986; Daughety 1990), and/or a larger num-ber of players (Sherali 1984; Heywood and McGinty 2008; Watt 2002). The overall sequential structure of a Stackelberg model is usually called a hierarchy (Boyer and Moreaux 1986; Watt 2002). In this context, a puzzling result has often been noticed: despite the hierarchical structure, firms behave as Cournotian oligopolists on the resid-ual demand. In the specific case of a single firm per stage, the standard Stackelberg duopoly leader produces exactly the same quantity as a monopolist. More generally, for a larger number of firms per stage and a homogeneous product, the strategies of leaders at any stage depend neither on the number of followers who play after, nor on the number of remaining stages. While this property is a striking feature of the Stack-elberg literature, neither its foundations, nor its implications have been fully formally studied.[1]

In this paper, we go a step further by investigating the conditions under which such a property holds. To do this, we determine the conditions under which a Stackelberg leader firm behaves as a Cournotian oligopolist. We then consider the maximiza-tion program of a Cournot oligopolist and we analyze the assumptions under which the program of a Stackelberg leader is equivalent. While the literature only focused on the equivalence between the equilibrium strategies, our approach allows us to study the impact on profit functions of alternative assumptions on market demand, costs and the hierarchical structure.[2] The equivalence between the best response functions and the equilibrium strategies follows.

We show that the property holds whatever the number of cohorts and firms in the economy, as long as we conjointly assume the three following assumptions: a linear demand, constant and identical marginal costs. These three assumptions, unlike the structure and size of the economy, are critical features. Assuming the linearity for market demand and costs is not sufficient since individual marginal costs have also to be equal for all firms. We fill a gap in the existing literature by showing that these assumptions are not only sufficient but also necessary conditions for a T -stage hierarchical oligopoly model to reduce to a succession of Cournot games. Any departure from one of these assumptions rules out the equivalence between the Cournot and the Stackelberg strategies. We display several examples to illustrate our results.

The paper is organized as follows. In Sect. 2, we present the hierarchical model, state the assumptions and provide a graphical interpretation of the property. The next two sections deal with the formal proof that the hierarchical model reduces to a succession of Cournot games. Section 3 considers sufficient conditions, while Sect. 4 deals with necessary conditions. In Sect. 5, we provide an illustration with three cohorts. Section 6 concludes.

[1] Pal and Sarkar (2001) and Lafay (2010), however, analyze the robustness of the property by investigating the impact of cost differences on the equilibrium strategies.

[2] Our analysis then does not restrict to equilibrium strategies.

2 The hierarchical model under linearity assumptions

2.1 General framework

Consider one homogeneous good produced by n firms which oligopolistically compete in a hierarchical framework. There are T stages of decisions indexed by t, $t \in [\![1, T]\!]$. Each stage embodies one cohort and is associated with a level of decision. The whole set of cohorts represents a hierarchy. Cohort t is populated by n_t firms, with $\sum_t n_t = n$. The distribution of the firms within each cohort is assumed to be observable and exogenous.[3] This latter assumption notably implies that position of firms and timing of moves are given.[4]

A firm i which belongs to cohort t has to decide strategically (simultaneously with firms of the same cohort, and sequentially among the hierarchy) its level of output denoted by x_t^i. The aggregate output of cohort t is denoted $X_t = \sum_{i=1}^{n_t} x_t^i$, where x_t^i stands for firm i's output within cohort t. In addition, $X_t^{-i} = \sum_{-i} x_t^{-i}$ denotes the production of all firms belonging to cohort t but i.

The n_t firms which belong to cohort t, behave as followers with respect to all firms of cohort τ, $\tau \in [\![1, t-1]\!]$, whose strategies are taken as given. However, they behave as Stackelberg leaders toward all firms of cohort τ, $\tau \in [\![t+1, T]\!]$. They consider the best-response functions of all firms belonging to these cohorts as functions of their strategies.

The inverse market demand function for the homogeneous good specifies the market price p as a function of aggregate output X, and is denoted by $p(X)$. We assume that $p(.)$ is continuous and twice differentiable, with $\frac{dp(X)}{dX} < 0$. For the equilibrium aggregate output X^*, the equilibrium price is unique and equal to $p(X^*)$.

The cost function of any firm i which belongs to cohort t is denoted by $\phi_t(.)$. It is a continuous and twice differentiable function with $\frac{d\phi_t(x_t^i)}{dx_t^i} > 0$ and $\frac{d^2\phi_t(x_t^i)}{d(x_t^i)^2} \geq 0$.

2.2 The linearity assumptions

In this section, we assume

- an inverse market demand function which is linear

$$p(X) = a - bX, \tag{H1}$$

  where $a,\ b > 0$ and $X = \sum_{t=1}^{T} X_t$,
- constant marginal costs

$$\phi_t(x_t^i) = c_t x_t^i, \quad i = 1, \ldots, n_t \qquad \text{and} \quad t = 1, \ldots, T \tag{H2}$$

[3] The standard Stackelberg duopoly prevails when $T = 2$ and $n_1 = n_2 = 1$.

[4] We therefore do not question the way a specific firm could or should become a leader (see Anderson and Engers 1992; Amir and Grilo 1999; Matsumura 1999).

- and identical marginal costs

$$c_t = c, \quad t = 1, \ldots, T. \tag{H3}$$

These three assumptions are standard in the literature on oligopoly analysis (see Daughety 1990; Carlton and Perloff 1994; Vives 1999, among others).

The puzzling result of the hierarchical model stressed in the introduction has been regularly stated in the literature (e.g. Boyer and Moreaux 1986; Watt 2002). It can be captured by the property below:

Property 1 *In the Stackelberg linear economy with an homogeneous product, firms in cohort t, $t \in [\![1, T]\!]$, behave as Cournotian oligopolists on the residual demand left by firms of cohorts $\tau (1 \leq \tau < t)$.*

While this property has been quoted in the standard literature, it has been understood in a very restrictive way: the equilibrium strategies of firms in the hierarchical model coincide with those of a multistage Cournot model. In this paper, we enrich the meaning and implications of Property 1, presented as the equivalence of the profit functions in the two models *up to a linear transformation*. This implies the equivalence of the reaction functions in both models and thereafter of the equilibrium strategies (which is then a consequence rather than a definition of the Cournotian behavior).

Property 1 is usually observed under assumptions (H1)–(H3). We intend to investigate the sufficiency and necessity of these assumptions in the occurrence of this property as it has been defined above. Sufficiency has been considered in the literature for the restrictive definition of Property 1 but can be formally established in our enlarged framework. Necessity of these assumptions has never been studied.

2.3 Implications: a graphical interpretation

In this subsection, we illustrate Property 1 by a graphical approach, which is based on demand and costs rather than on best response functions. This property will be formally studied in the next section.

Consider two successive stages, say $t - 1$ and t, with $t \neq T$. At stage $t - 1$, the aggregate output resulting from the individual productions of firms of cohorts 1 to $t - 1$ is given by $\sum_{\tau=1}^{t-1} X_\tau$. No particular value is given to X_τ, except that it must generate a non-negative profit and is decided once for all. Under Property 1, the aggregate production of cohort t is computed as if it were the last cohort. In this case, the aggregate output of cohort t (for $\sum_{\tau=1}^{t-1} X_\tau$) would establish to $X_t = \sum_{\tau=1}^{t} X_\tau - \sum_{\tau=1}^{t-1} X_\tau$. The subsequent price $p\left(\sum_{\tau=1}^{t} X_\tau\right)$ is the market price that would be effective if there were only t cohorts in the hierarchy. However, since cohort t is not the last cohort, the (unique) equilibrium price of the economy is lower and equal to $p(\sum_{\tau=1}^{T} X_\tau)$. Nevertheless, the aggregate output of cohort t remains unchanged when the price falls down to $p\left(\sum_{\tau=1}^{T} X_\tau\right)$.

We assume that any leading cohort $\tau < t$ expects firms of cohort t (or more) to act symmetrically while firms of cohort t maximize their profit for any given quantity

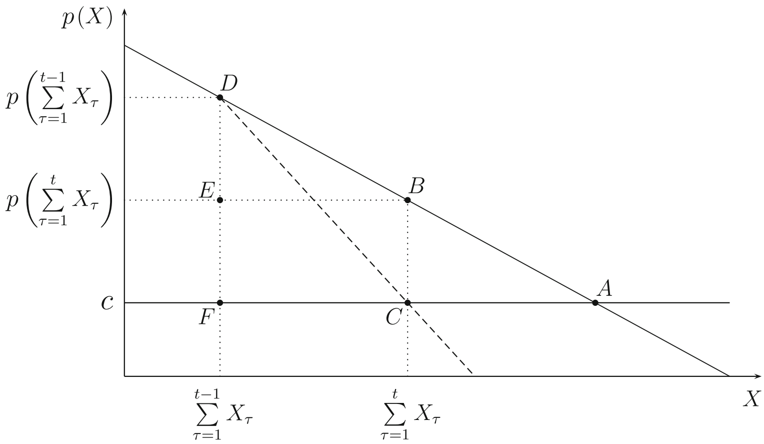

Fig. 1 Graphical representation of (H1)–(H3)

$\sum_{\tau=1}^{t-1} X_\tau$ produced by their predecessors. The overall situation for cohort t is depicted in Fig. 1.

In this figure, we illustrate the behavior of cohort-t firms when acting as Cournotian oligopolists on the residual demand left by firms of cohort $\tau < t$, as if they did not take into consideration firms playing after. In other words, Property 1 implies that leaders choose their optimal quantities so as to equalize the Cournotian marginal revenue of the cohort (dashed line) to the marginal cost (full line). As an immediate consequence, strategies of the leaders at any stage depend neither on the number of followers who play after nor on the remaining number of stages. This consequence corresponds to the standard result about the linear model usually expressed in the literature: given that the strategies of leaders do not depend on the number of followers, leaders behave as if they played a Cournot game on residual demand.

3 Sufficiency

3.1 The linearity assumptions as sufficiency conditions

Demonstrating Property 1 requires to exhibit the link between leaders and followers' profits.

Lemma 1 *Let* $\gamma_t \equiv \prod_{\tau=t+1}^{T} \frac{1}{1+n_\tau}$ *be the leader's markup discount factor. Under* (H1)–(H3)*, the markup earned by a cohort t firm,* $t < T$*, in a* T*-cohort economy is a constant share* $\gamma_t < 1$ *of the markup it earns in a* t*-cohort economy for any given vector of outputs* $(X_1, \ldots, X_{t-1})$ *produced by the previous cohorts:*

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c = \gamma_t \left[p\left(\sum_{\tau=1}^{t} X_\tau\right) - c \right] \quad \textit{for } t < T. \tag{1}$$

Proof See Appendix A.

Notice that under conditions H2 and H3, the markup is always equal across cohorts. The discount factor γ_t differs from one cohort to another and measures the impact on market power of the number of followers. More specifically, it represents the reduction of a leader's markup due to the presence of the additional cohorts $t+1$ to T. It affects less intensively the market power of the last cohorts in the sequence since they face a reduced number of followers. Market power shrinks as t tends to infinity.

The existence of cohort τ equally impacts by a coefficient $1/(1+n_\tau)$ the markup expected by a leader t ($t < \tau$) in a t-stage economy, whatever the quantities produced by the first t cohorts (this results directly from H1–H3).

Corollary 1 *For any strategy x_t^i, the profit obtained by a cohort—t firm in the sequential T-stage structure is a constant share of the profit of a t-stage economy* (*where cohort t would be the last cohort*)*:*

$$\pi_t^i(x_t^i) = \left[p\left(\sum_{\tau=1}^{T} X_\tau\right) - c\right] x_t^i = \gamma_t \left[p\left(\sum_{\tau=1}^{t} X_\tau\right) - c\right] x_t^i. \qquad (2)$$

Proof This corollary directly results from Lemma 1.

In other words, each cohort can behave *as if* there were no following cohorts since it earns a constant share of the profit realized in an oligopoly structure market where it represents the last cohort, whatever the aggregate output $\sum_{\tau=1}^{t-1} X_\tau$ produced by the leaders. Provided that cohort-t firms maximize their profit for any vector of strategies $(X_1, \ldots, X_{t-1})$, cohort-τ leaders ($\tau < t$) act as Cournot players who choose simultaneously and assume their direct rivals' output as given, ignoring the following cohorts, that only discount the value of their profits without changing the nature of the maximization program.

In the next sections, the right-hand side of Eq. (2) divided by γ_t will be called the *myopic profit* (see Appendix A). By contrast with the sequential structure of our economy (left-hand side of Eq. (2)), the associated *myopic program* would result from the behavior of a firm that would observe the strategies of the leaders and would ignore the supplies of its followers.

Lemma 2 *Let $\eta_{t-h,t} \equiv \prod_{\tau=t-h+1}^{t} \frac{1}{1+n_\tau}$ be the follower's output discount factor. Under* (H1)–(H3)*, the output of a firm i in cohort $t \leq T$ can be expressed as a share of the output produced by a firm playing previously and belonging to cohort $t-h$ for $h \in [\![1, t-1]\!]$, that is:*

$$x_t = \eta_{t-h,t} x_{t-h}.$$

Proof See Appendix B.

The follower's output discount represents the decrease in optimal quantities for a follower resulting from a contraction of the residual demand when playing latter in the hierarchy. It is a share of cohort $t-h$'s output, whether optimal or not, which is optimal for cohort t to produce. For any follower, this share decreases when going further in the sequence. Its value depends negatively on the number of leading cohorts and on the number of corresponding players (because residual demand decreases with respect to this parameter).

From the previous lemmas, the following proposition can be stated:

Proposition 1 *Assumptions* (H1)–(H3) *are sufficient conditions for Property 1 to hold: the T-stage Stackelberg linear economy reduces to a collection of Cournot games in which firms compete oligopolistically on residual demand.*

Proof The proposition directly ensues from Lemmas 1 and 2.

3.2 Characterization of the linear hierarchical equilibrium

Maximizing the *sequential profit* (left-hand side of Eq. (2)) is tantamount to maximize the *myopic profit* (right-hand side of Eq. (2)) since γ_t is a constant term. In the linear economy, strategies of firms do not depend on the number of firms playing after, which equally impact the profit associated to each strategy. As a consequence the optimal strategies and the equilibrium strategies remain unchanged whatever the number of stages and the number of followers.[5]

The literature only covers the similarity of the equilibrium strategies in both the T-stage Stackelberg linear model and the collection of staggered static problems but does not provide any explanation for this coincidence (see Boyer and Moreaux 1986; Anderson and Engers 1992; Watt 2002).

Corollary 2 *The equilibrium strategy of cohort-*1 *firms may be obtained either from the sequential profit maximization or from the myopic profit maximization:*

$$x_1 = \frac{1}{1+n_1}\frac{a-c}{b} \equiv \eta_1 X^*,$$

where $X^* = (a-c)/b$ is equal to the perfect competition aggregate output. We then deduce the equilibrium strategy of any firm i in cohort t, $t \in [\![1, T]\!]$:

$$x_t = \frac{\eta_{1,t}}{1+n_1}\frac{a-c}{b} \equiv \eta_t X^* \quad \textit{with } \eta_t \equiv \prod_{\tau=1}^{t} \frac{1}{1+n_\tau}.$$

Notice that in the equilibrium, each firm of cohort t produces a share η_t of the perfect competition equilibrium output.

[5] Property 1 still holds in the presence of static conjectural variations (see Julien and Musy 2011).

Corollary 3 *The equilibrium price and profits are given by:*

$$p = c + (a-c)\prod_{\tau=1}^{T}\frac{1}{1+n_\tau},$$

$$\pi_t^i = \frac{(a-c)^2}{b}\prod_{\tau=1}^{t}\frac{1}{(1+n_\tau)^2}\prod_{\tau=t+1}^{T}\frac{1}{1+n_\tau}, \quad t = 1,\ldots,T-1$$

$$\pi_T^i = \frac{(a-c)^2}{b}\prod_{\tau=1}^{T}\frac{1}{(1+n_\tau)^2}.$$

De Quinto and Watt (2003) use a term similar to η_t to analyze welfare through market power and mergers.

Remark 1 Leaders of any cohort t, $t \in [\![1, T-1]\!]$, do not ignore the true number of followers in cohort $\tau > t$ or do not form incorrect estimates about it.

Remark 2 The leader's markup discount factor is equivalent to a taxation of profits. The tax rate $\tau_t = 1 - \gamma_t$ differs from one firm to another.

4 Beyond the sufficiency of the linearity assumptions

This section establishes the following proposition and provides some examples.

Proposition 2 *Assumptions* (H1)–(H3) *are necessary conditions for Property 1 to hold when costs are strictly positive: the T-stage Stackelberg linear economy reduces to a collection of Cournot games in which firms compete oligopolistically on the residual demands.*

In the literature, neither the necessity of these conditions nor the fact that the leaders effectively behave as Cournotian oligopolists have ever been demonstrated. The only results feature that the equilibrium strategies in a T-stage model coincides with the equilibrium strategies that would be obtained with staggered static problems where firms either compete monopolistically (for $n_t = 1$) or oligopolistically on the residual demand.

4.1 Necessity of assumption (H1)

We have to show that the linearity of the inverse demand function is a necessary condition for firms to behave as Cournotian oligopolists in each cohort, provided that the constant marginal cost is different from zero.

4.1.1 Proof

Under assumptions (H2) and (H3), firms of cohort $T-1$ act as Cournotian competitors on residual demand if and only if:

$$\left[p\left(\sum_{\tau=1}^{T-1} X_\tau + X_T\right) - c\right] x_t = \gamma_t \left[p\left(\sum_{\tau=1}^{T-1} X_\tau\right) - c\right] x_t, \tag{3}$$

for a coefficient $\gamma_t \equiv \prod_{\tau=t+1}^{T} \frac{1}{1+n_\tau} > 0$, or equivalently:

$$p\left(\sum_{\tau=1}^{T-1} X_\tau + X_T\right) - c = \gamma_t \left[p\left(\sum_{\tau=1}^{T-1} X_\tau\right) - c\right]. \tag{4}$$

The conditions on the inverse demand function under which the equality in (4) is satisfied must be established. For a strictly positive marginal cost, the only inverse demand function satisfying (4) is linear. The proof is derived in two steps. We first need to demonstrate that Eq. (4) requires the reaction function of the last cohort to be linear. Second, we check that this property occurs only if the inverse demand function is linear.

Denote by g the function defined by $g(.) = p(.) - c$ and by f the reaction function of the last cohort: $X_T = f(\sum_{\tau=1}^{T-1} X_\tau)$. Let $\mathcal{O}_g$ and $\mathcal{O}_f$ be the orders of the terms $g(\sum_{\tau=1}^{T-1} X_\tau)$ and $f(\sum_{\tau=1}^{T-1} X_\tau)$.

Since γ_t is a constant, it is an order-0 term and $\gamma_t g(\sum_{\tau=1}^{T-1} X_\tau)$ has the same order as $g(\sum_{\tau=1}^{T-1} X_\tau)$, that is $\mathcal{O}_g$. The left-hand side of Eq. (4) is a composed function. The term $g[\sum_{\tau=1}^{T-1} X_\tau + f(\sum_{\tau=1}^{T-1} X_\tau)]$ is a term of order $\mathcal{O}_g \mathcal{O}_f$ for $\mathcal{O}_f \neq 0$ and of order $\mathcal{O}_g$ for $\mathcal{O}_f = 0$. In the former case (with $\mathcal{O}_g \neq 0$), the equality in (4) requires that $\mathcal{O}_f = 1$, which means that f be a linear function.[6]

Consequently, the aggregate reaction function of cohort T must be linear in $\sum_{\tau=1}^{T-1} X_\tau$. Formally:[7]

$$X_T = \alpha_1 + \alpha_2 \sum_{\tau=1}^{T-1} X_\tau, \tag{5}$$

where $\alpha_1 > 0$ is the aggregate output of cohort T when the other cohorts do not produce, and $-1 < \alpha_2 \leq 0$ is the slope of the reaction function of cohort T. The values of those parameters are determined on the ground of economic plausibility. Strategies are said to be substitute for $\alpha_2 < 0$ and independent for $\alpha_2 = 0$.

Adding $\alpha_2 X_T$ to each side of the equation above and rearranging yields:

$$X_T = \frac{\alpha_1}{1+\alpha_2} + \frac{\alpha_2}{1+\alpha_2} \sum_{\tau=1}^{T} X_\tau. \tag{6}$$

[6] The case $\mathcal{O}_g = 0$ means that price is constant, that is independent of aggregate output.

[7] The case $\mathcal{O}_f = 0$ is a subcase of Eq. (5) for which $\alpha_2 = 0$.

Given Eq. (2), the first-order condition for profit maximization of firm i belonging to cohort T is:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) + \frac{dp}{dX}\left(\sum_{\tau=1}^{T} X_\tau\right) x_T^i - c = 0,\ i = 1, \ldots, n_T. \tag{7}$$

The n_T equations in (7) implicitly define the reaction functions of cohort-T firms. Summing those equations for all i, $i = 1, \ldots, n_T$, leads to the reaction function of cohort T. This function must be consistent with the form determined in (6).

Summing Eqs. (7) for all firms i, $i = 1, \ldots, n_T$, that belong to cohort T and using Eq. (6) to substitute for X_T yields:

$$-\frac{p(X) - c}{dp/dX} = \mu_1 + \mu_2 X, \tag{8}$$

with: $\mu_1 = \dfrac{\alpha_1}{n_T(1+\alpha_2)} > 0$, $\mu_2 = \dfrac{\alpha_2}{n_T(1+\alpha_2)} \leq 0$ and $X = \sum_{\tau=1}^{T} X_\tau$.

Finding the condition(s) on the inverse demand function p under which Eq. (4) is satisfied requires to solve Eq. (8). By inverting and integrating both sides of Eq. (8), we get:

$$\int \frac{dp/dX}{p(X) - c}\, dX = -\frac{1}{\mu_2}\int \frac{\mu_2}{\mu_1 + \mu_2 X}\, dX, \tag{9}$$

whose solution is:

$$\begin{cases} \ln[p(X) - c] = -\frac{1}{\mu_2}\ln[\mu_3(\mu_1 + \mu_2 X)] & \text{for } \mu_2 \neq 0 \\ \ln[p(X) - c] = -\frac{1}{\mu_1}X + \ln \mu_4 & \text{for } \mu_2 = 0, \end{cases} \tag{10}$$

where $\mu_3 > 0$ and $\mu_4 > 0$ stand for the constants of integration.

The inverse demand function is finally obtained by identification. There is no economic rationale to consider that the inverse demand function is defined with respect to c. Thus, for $c \neq 0$, all terms in c must disappear from Eq. (10). Two cases must be considered.

Case 1: $c = 0$

In that case, terms in c disappear from Eq. (10) and the values of the parameters μ_1 to μ_4 may be freely chosen to determine the inverse demand function.

For $\mu_2 \neq 0$, Eq. (10) becomes:

$$p(X) = [\mu_3(\mu_1 + \mu_2 X)]^{-1/\mu_2} \equiv [\alpha - \beta X]^{\rho} \text{ for any } \alpha, \beta, \rho > 0.$$

A specific case is the linear demand function with $\rho = 1$. It is worth noting however that for $\rho \neq 1$, another class of non-linear functions is compatible with the property that firms behave a la Cournot on residual demand when $c = 0$.[8]

For $\mu_2 = 0$, Eq. (10) collapses to:

$$p(X) = \mu_4 e^{-X/\mu_1} \equiv \alpha e^{-\rho X} \quad \text{for any } \alpha, \rho > 0$$

which defines a last class of non-linear demand functions.

Case 2: $c \neq 0$

For $\mu_2 = 0$, Eq. (10) collapses to:

$$p(X) = \mu_4 e^{-X/\mu_1} + c.$$

Such a solution implies that the inverse demand function directly depends on the marginal cost c, which has no economic meaning. So this case can be ruled out.

For $\mu_2 \neq 0$, Eq. (10) becomes:

$$p(X) - c = [\mu_3(\mu_1 + \mu_2 X)]^{-1/\mu_2},$$

We isolate the class of demand functions independent of c. Therefore, for sake of tractability, we use a first-order Taylor expansion around $\tilde{X}$ to approximate the different inverse demand functions.

$$\begin{aligned} p(X) - c &\simeq p(\tilde{X}) - c - \mu_3[\mu_3(\mu_1 + \mu_2 X)]^{-1/\mu_2 - 1}(X - \tilde{X}) \\ p(X) &\simeq p(\tilde{X}) - \mu_3[\mu_3(\mu_1 + \mu_2 \tilde{X})]^{-1/\mu_2 - 1}(X - \tilde{X}). \end{aligned} \tag{11}$$

In the above equation, only parameters μ_1 and μ_2 may depend on c (μ_3 being a constant of integration). We now show that μ_1 necessarily depends on c for $\alpha_1 \neq \kappa(1 + \alpha_2)$, with $\kappa \in \mathbb{R}_{++}$, a case that will be dealt with later. Notice first that α_1 must depend on c since Eq. (7) must be satisfied, notably for $\sum_{\tau=1}^{T-1} X_\tau = 0$. Since p is independent of c, x_T (and then X_T) must depend on c. For $\sum_{\tau=1}^{T-1} X_\tau = 0$, recall that $X_T = \alpha_1$. According to (8), parameter μ_1 also depends on c provided $\alpha_1 \neq \kappa(1+\alpha_2)$, with $\kappa \in \mathbb{R}_{++}$, which implies that μ_1 must disappear from the previous equation.

When μ_2 does not embody terms in c, the only possible value for μ_1 to vanish is $\mu_2 = -1$. In that case: $p(X) \simeq p(\tilde{X}) - \mu_3(X - \tilde{X})$, where the value for μ_3 can be freely chosen. Set, for instance, $\mu_3 = \beta > 0$. Then, the inverse demand function is:

$$p(X) = \alpha - \beta X \quad \text{for any } \alpha, \beta > 0, \tag{12}$$

where α and β are parameters independent of c, and $\mu_1 = (\alpha - c)/\beta > 0$ according to Eq. (8). Equation (12) corresponds to assumption (H1). It satisfies Property 1 for any $(\alpha, \beta) \in \mathbb{R}^2_{++}$.

When μ_2 includes terms in c, the right-hand side of the Taylor expansion is independent of c if $\mu_1 + \mu_2 \tilde{X} = 0$ or $\mu_3(\mu_1 + \mu_2 \tilde{X}) = 1$. From (8), the first case corresponds

[8] This result was pointed out by Anderson and Engers (1992) in footnote 5 on page 129.

to perfect competition: $p = c$. The second means that price is constant for any value of the aggregate output: $p = c + 1$. In both cases, the inverse demand function is dependent of c. These cases can be ruled out.

Thus, the linear demand function is the only class of functions consistent with the property that firms act as Cournotian oligopolist on residual demand when $c \neq 0$.

4.1.2 Example for assumption (H1)

Consider $T = 2$ and $n_1 = n_2 = 1$. The inverse market demand function is given by $P(X) = e^{-X^2}$. One can derive the best response function of the follower $x_2(x_1) = \left(\sqrt{x_1^2 + 2} - x_1\right)/2$. The equilibrium strategy of the leader is then $\tilde{x}_1 = \sqrt{\sqrt{2} - 1}$. These strategies differ from the equilibrium outputs that would be obtained if there were only one monopolistic firm, i.e. $x^m = \sqrt{2}/2$.

4.2 Necessity of assumption (H3)

4.2.1 Proof

Under assumptions (H1) and (H2), the reaction function of a cohort-T firm (assuming symmetry of strategies across cohort-T) can now be explicitly defined:

$$x_T = \frac{a - c_T - b \sum_{\tau=1}^{T-1} X_\tau}{b(1 + n_T)}, \tag{13}$$

and the profit function of cohort-$(T-1)$ firms is:

$$\pi_{T-1}(x_{T-1}) = \frac{1}{1 + n_T} \left[\bar{a}_{T-1} - bX_{T-1} - c_{T-1} - n_T(c_{T-1} - c_T)\right] x_{T-1}, \tag{14}$$

where $\bar{a}_t \equiv a - b \sum\limits_{\tau=1}^{t} X_\tau$ is the residual demand left by the first t cohorts.

For cohort-$(T-1)$ firms to behave as Cournotian oligopolists, the profit function must be such that:

$$\pi_{T-1}(x_{T-1}) = \frac{1}{1 + n_T} \tilde{\pi}_{T-1}(x_{T-1}). \tag{15}$$

This is the case if and only if:

$$c_{T-1} = c_T.$$

Using backward induction, it can be proved that for cohort-t firms, $t < T - 1$, to behave as Cournotian oligopolists, the following property must hold:

$$c_1 = c_2 = \cdots = c_{T-1} = c_T. \tag{16}$$

Thus, when the inverse demand function is linear and cost functions are constant, a necessary condition for firms of any cohort to behave as Cournotian competitors on residual demand requires assumption (H3).

4.2.2 Example for assumption (H3)

Consider $T = 2$, n_1 leaders and n_2 followers. The inverse market demand function is $P(X) = a - bX$. Assume $\phi_t(x_t^i) = c_t x_t^i$, $t = 1, 2$. One easily gets $\tilde{x}_1 = \frac{(a-c_1)+(c_2-c_1)n_2}{b(1+n_1)}$. For $c_1 \neq c_2$, the equilibrium strategy of leaders depends on the number of followers. It is no longer the case for $c_1 = c_2$ since $\tilde{x}_1 = \frac{a-c}{b(1+n_1)}$.

4.3 Necessity of assumption (H2)

4.3.1 Proof

Assume that assumptions (H1) and (H3) are satisfied. The firm j of cohort t then determines x_t^j such that it maximizes its myopic profit function $\tilde{\pi}_t^j$:

$$\tilde{\pi}_t^j(x_t^j) = \left[\bar{a}_{t-1} - b\sum_{i=1}^{n_t} x_t^i\right] x_t^j - \phi(x_t^j), \tag{17}$$

where ϕ is a non-decreasing function and is identical across firms.

The reaction function of cohort-t firms, when behaving symmetrically, is implicitly defined by:

$$\bar{a}_{t-1} - b(1+n_t)x_t - \frac{d\phi}{dx_t}(x_t) = 0 \quad \text{for any } t > 1. \tag{18}$$

When maximizing its profit, a cohort-τ firm ($\tau < t$) substitutes to X_t the aggregate reaction function of cohort t. Strategy of a cohort-τ firm is then independent of n_t provided X_t does not contain any n_t term. If so x_t is a function of n_t, i.e. $x_t = X_t/(1+n_t)$.

The aggregate reaction function of cohort t can be rewritten as:

$$X_t = \frac{\bar{a}_{t-1} - \frac{d\phi}{dx_t}(x_t)}{b} \quad \text{for any } t > 1. \tag{19}$$

It does not depend on n_t provided $\frac{d\phi_t}{dx_t}(x_t)$ does not contain any n_t term. Since x_t is a function of n_t, the marginal cost is independent of x_t, that is constant:

$$\phi(x_t) = cx_t \quad \text{for any} t \in [\![1, T]\!]. \tag{20}$$

4.3.2 Example for assumption (H2)

Consider $T = 2$, n_1 leaders and n_2 followers. The inverse market function is $P(X) = a - bX$. Assume $\phi_t(x_t^i) = c(x_t^i)^2/2$ for $i = 1, \ldots, n_t$ and $t = 1, 2$. One gets $\tilde{x}_1 = \frac{a}{b(1+n_1)+c\left(1+\frac{b}{b+c}n_2\right)}$, so the equilibrium strategy of any leader depends on the number of followers. When firms behave *a la* Cournot in each stage, one has $\hat{x}_1 = \frac{a}{b(1+n_1)+c} \neq \tilde{x}_1$.

5 An illustration with three cohorts

5.1 A three-stage Stackelberg economy

Consider as an illustration an economy with $T = 3$ under assumptions (H1)–(H3). In this three-stage game, cohorts 1 to 3 are populated by respectively n_1, n_2 and n_3 firms. Cohort 1 plays first but we solve the game by backward induction, starting from the last cohort.

The program of firm k, $k = 1, \ldots, n_3$, which belongs to cohort 3 may be written:

$$Arg \max_{x_3^k} \pi_3^k = \left[a - b\left(X_1 + X_2 + x_3^k + X_3^{-k}\right) - c\right] x_3^k,$$

that is:

$$Arg \max_{x_3^k} \pi_3^k = \left[a - b\sum_{i=1}^{n_1} x_1^i - b\sum_{j=1}^{n_2} x_2^j - b(x_3^k + X_3^{-k}) - c\right] x_3^k, \tag{21}$$

where x_1^i (resp. x_2^j) is the strategic supply of firm i (resp. j) in cohort 1 (resp. 2), observed by firm k when choosing its own strategy. Each firm k considers as given the strategies of its $n_3 - 1$ competitors in cohort 3. In the symmetric equilibrium: $x_3^k = x_3^{-k}$ and $X_3^{-k} = (n_3 - 1)x_3^{-k}$. The first-order condition leads to the reaction function of any firm k:

$$x_3^k(X_1, X_2) = \frac{1}{1 + n_3}\frac{a - c}{b} - \frac{X_1 + X_2}{1 + n_3}. \tag{22}$$

When choosing its own strategy, each firm j belonging to cohort 2 observes the strategic supplies of the n_1 leaders and considers as given the strategies of its $n_2 - 1$ competitors. It also takes into account the reaction function of cohort-3 firms. Then, the program of the jth cohort-2 firm may thus be written:

$$Arg \max_{x_2^j} \pi_2^j = \left[a - b\left(X_1 + x_2^j + X_2^{-j} + \sum_{k=1}^{n_3} x_3^k(X_1, X_2)\right) - c\right] x_2^j,$$

that is:

$$\underset{x_2^j}{Arg \max}\, \pi_2^j = \frac{1}{1+n_3}\left[a - b\sum_{i=1}^{n_1} x_1^i - b(x_2^j + X_2^{-j}) - c\right] x_2^j. \tag{23}$$

In the symmetric equilibrium: $x_2^j = x_2^{-j}$ and $X_2^{-j} = (n_2 - 1)x_2^{-j}$. The first-order condition leads to the reaction function of any firm j:

$$x_2^j(X_1) = \frac{1}{1+n_2}\frac{a-c}{b} - \frac{X_1}{1+n_2} \quad j = 1, \ldots, n_2 \tag{24}$$

Finally, firm i belonging to cohort 1 takes into account the reaction functions of all its followers and considers as given the strategies of its $n_1 - 1$ competitors. Its program may be written:

$$\underset{x_1^i}{Arg \max}\, \pi_1^i = \left[a - b\left(x_1^i + X_1^{-i} + \sum_{j=1}^{n_2} x_2^j(X_1) + \sum_{k=1}^{n_3} x_3^k(X_1, X_2)\right) - c\right] x_1^i,$$

that is:

$$\underset{x_1^i}{Arg \max}\, \pi_1^i = \frac{1}{(1+n_2)(1+n_3)}[a - b(x_1^i + X_1^{-i}) - c]x_1^i. \tag{25}$$

In the symmetric equilibrium, one gets the equilibrium strategy of leader i:

$$\tilde{x}_1^i = \frac{1}{1+n_1}\frac{a-c}{b} \quad i = 1, \ldots, n_1. \tag{26}$$

We deduce the equilibrium outputs of cohort-1 and cohort-2 firms:

$$\tilde{x}_2^j = \frac{1}{(1+n_1)(1+n_2)}\frac{a-c}{b} \quad j = 1, \ldots, n_2 \tag{27}$$

$$\tilde{x}_3^k = \frac{1}{(1+n_1)(1+n_2)(1+n_3)}\frac{a-c}{b} \quad k = 1, \ldots, n_3 \tag{28}$$

As specified in Property 1, the equilibrium strategy of any leader in the first stage depends neither on the number of followers who play after, nor on the number of remaining stages. This property also holds for firms in stage 2.

The corresponding aggregate outputs in the three stages are:

$$\begin{aligned}\tilde{X}_1 &= \frac{n_1}{1+n_1}\frac{a-c}{b}\\ \tilde{X}_2 &= \frac{n_2}{(1+n_1)(1+n_2)}\frac{a-c}{b}\\ \tilde{X}_3 &= \frac{n_3}{(1+n_1)(1+n_2)(1+n_3)}\frac{a-c}{b}\end{aligned} \tag{29}$$

The corresponding equilibrium market price and profits are given by:

$$\tilde{p} = c + \frac{a-c}{(1+n_1)(1+n_2)(1+n_3)}, \tag{30}$$

and

$$\begin{aligned}\tilde{\pi}_1^i &= \frac{(a-c)^2}{b}\frac{1}{(1+n_1)^2(1+n_2)(1+n_3)}\\ \tilde{\pi}_2^j &= \frac{(a-c)^2}{b}\frac{1}{(1+n_1)^2(1+n_2)^2(1+n_3)}\\ \tilde{\pi}_3^k &= \frac{(a-c)^2}{b}\frac{1}{(1+n_1)^2(1+n_2)^2(1+n_3)^2}\end{aligned} \tag{31}$$

So the equilibrium values coincide with those given in Corollary 3.

5.2 A three-stage Cournot economy

Consider now a three-stage Cournotian economy.

In stage 1, any firm i, $i = 1, \dots, n_1$, solves:

$$\underset{x_1^i}{Arg\max}\ \pi_1^i = [a - b(x_1^i + X_1^{-i}) - c]x_1^i. \tag{32}$$

Notice that this program is identical to the program of Eq. (25) up to a factor $\eta_{1,3} = \frac{1}{(1+n_2)(1+n_3)}$. The equilibrium strategy $\hat{x}_1^i$ of firm i must then be identical to the solution of Eq. (25):

$$\hat{x}_1^i = \frac{1}{1+n_1}\frac{a-c}{b} = \tilde{x}_1^i, \quad i = 1, \dots, n_1. \tag{33}$$

In stage 2, any firm j, $j = 1, \dots n_2$, determines its supply from the residual demand left by cohort-1 firms:

$$\underset{x_2^j}{Arg\max}\ \pi_2^j = \left[a - b\sum_{i=1}^{n_1}\hat{x}_1^i - b(x_2^j + X_2^{-j}) - c\right]x_2^j,$$

that is:

$$Arg \max_{x_2^j} \pi_2^j = \frac{1}{1+n_1}\left[a - b\sum_{i=1}^{n_1}\hat{x}_1^i - b(x_2^j + X_2^{-j}) - c\right]x_2^j. \tag{34}$$

Provided the aggregate outputs of cohort 1 in both economies (whether optimal or not) are equal, that is $\hat{X}_1 = \tilde{X}_1$, this program is identical to the program of Eq. (23) up to a factor $\eta_{2,3} = \frac{1}{1+n_3}$. The equilibrium strategy $\hat{x}_2^j$ of firm j must then be identical to the solution of Eq. (23):

$$\hat{x}_2^j = \frac{1}{(1+n_1)(1+n_2)}\frac{a-c}{b} = \tilde{x}_2^j, \quad j = 1, \ldots, n_2. \tag{35}$$

Finally, any firm k, $k = 1, \ldots, n_3$, belonging to cohort 3 determines its supply from the residual demand left by cohort-1 and -2 firms:

$$Arg \max_{x_3^k} \pi_3^k = \left[a - b\sum_{i=1}^{n_1}\hat{x}_1^i - b\sum_{j=1}^{n_2}\hat{x}_2^j - b(x_3^k + X_3^{-k}) - c\right]x_3^k. \tag{36}$$

Provided the joint aggregate outputs of cohorts 1 and 2 in both economies (whether optimal or not) are equal, that is $\hat{X}_1 + \hat{X}_2 = \tilde{X}_1 + \tilde{X}_2$, this program is identical to the program of Eq. (21). The equilibrium strategy $\hat{x}_3^k$ of firm k must then be identical to the solution of Eq. (21):

$$\hat{x}_3^k = \frac{1}{(1+n_1)(1+n_2)(1+n_3)}\frac{a-c}{b} = \tilde{x}_3^k, \quad k = 1, \ldots, n_3. \tag{37}$$

The concordance of the Stackelberg and Cournotian maximization programs (up to a factor) and the resulting equilibrium strategies illustrates Property 1.

6 Conclusion

The paper explored a general hierarchical linear model in which firms compete in quantities. A key property has been studied: under the assumptions of linear demand, constant and identical marginal costs, each firm behaves at any stage as a Cournotian oligopolist on residual demand. We formally established that these assumptions are not only sufficient but also necessary for strictly positive marginal costs.

This property of the linear hierarchical T-stage model means that a firm in cohort t maximizes its profit when it chooses its output *as if* it were the last follower of the economy. The associated maximization program of the firm is equivalent up to a discount factor to the maximization program it would face in a similar t-stage hierarchical economy. The sooner a firm plays within the T-stage hierarchy, the more discounted its profit compared with the t-stage economy.

Acknowledgments We are grateful to C. Bidard for his comments and remarks on a previous version of the paper. Two anonymous referees are acknowledged for their remarks and suggestions. Remaining errors are ours.

Appendix A: Proof of Lemma 1

The proof is by backward induction and structured in three steps.

Step 1 property 1 is true for cohort $t = T - 1$ (with $T > 1$).

The inverse demand function faced by firms is defined by:

$$p(X) = a - bX \quad \text{with } X = \sum_{\tau=1}^{T} X_\tau, \tag{H1}$$

where $X_\tau = \sum_{i=1}^{n_\tau} x_\tau^i \geq 0$ is the aggregate production of cohort τ. For any quantity of output X_{T-1} produced by cohort $T - 1$, the resulting residual demand faced by followers of cohort T is:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) = \bar{a}_{T-1} - bX_T \quad \text{with } \bar{a}_t \equiv a - b\sum_{\tau=1}^{t} X_\tau, \tag{38}$$

where $\bar{a}_{T-1}$ is considered as given by followers. Geometrically (see Fig. 1), followers must select a couple (X, p) on the segment $[D, A]$.

When acting symmetrically, the associated marginal revenue of cohort-T firms (dashed line on Fig. 1) is defined by:[9]

$$R_m(X_T) = \bar{a}_{T-1} - b\frac{1 + n_T}{n_T} X_T. \tag{39}$$

Considering the following derivatives:

$$\frac{\partial R_m}{\partial X_T}(X_T) = \frac{DF}{CF} = -b\frac{1 + n_T}{n_T} \tag{40}$$

$$\frac{\partial p}{\partial X_T}(X_T) = \frac{DF}{AF} = -b, \tag{41}$$

it comes that:

$$CF = \frac{n_T}{1 + n_T} AF, \quad \text{or equivalently } AC = \frac{1}{1 + n_T} AF. \tag{42}$$

[9] This function is derived from the total revenue of a follower i: $RT(x_T^i) = [\bar{a}_{T-1} - b\sum_{k=1}^{n_T} x_T^k]x_T^i$. The symmetric behavior assumed for followers yields: $x_T^i = x_T$ for all $i \in [1, n_T]$ and $X_T = n_T x_T$.

Finally, applying Thales' theorem to triangles ABC and ADF leads to:

$$BC = \frac{1}{1+n_T} DF = EF. \tag{43}$$

Actually, EF is the markup of a leader after the entrance of the last cohort, while DF is the markup of a leader before the entrance of cohort T. Equation (43) can be rewritten as:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c = \frac{1}{1+n_T}\left[p\left(\sum_{\tau=1}^{T-1} X_\tau\right) - c\right]. \tag{44}$$

Step 2 assume property (1) is true for any cohort $t = T - h$ $(1 \leq h \leq T-2)$ then it is true for cohort $T - h - 1$.

If property (1) holds for cohort $T - h$ then:

$$\left[p\left(\sum_{\tau=1}^{T} X_\tau\right) - c\right] x^i_{T-h} = \gamma_{T-h}\left[p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c\right] x^i_{T-h}, \tag{45}$$

with $\gamma_{T-h} \equiv \prod_{\tau=T-h+1}^{T} \frac{1}{1+n_\tau}$.

Thus, maximizing firm i's profit is tantamount to maximize the myopic profit defined as follows:

$$\max_{x^i_{T-h}} \left[p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c\right] x^i_{T-h}. \tag{46}$$

When firms of cohort-$(T-h)$ act symmetrically, the myopic marginal revenue (dashed line) is defined by:[10]

$$\tilde{R}_m(X_{T-h}) = \bar{a}_{T-h-1} - b\frac{1+n_{T-h}}{n_{T-h}} X_{T-h}. \tag{47}$$

In the same way as in step 1, it can be shown that:

$$p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c = \frac{1}{1+n_{T-h}}\left[p\left(\sum_{\tau=1}^{T-h-1} X_\tau\right) - c\right]. \tag{48}$$

By assumption, the following property is satisfied:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c = \gamma_{T-h}\left[p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c\right]. \tag{49}$$

[10] The associated marginal revenue is: $R_m(X_{T-h}) = \tilde{R}_m(X_{T-h}) + (1 - \gamma_{T-h})c$.

We deduce from the two previous equations that:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c = \frac{\gamma_{T-h}}{1+n_{T-h}}\left[p\left(\sum_{\tau=1}^{T-h-1} X_\tau\right) - c\right] \tag{50}$$

$$= \gamma_{T-h-1}\left[p\left(\sum_{\tau=1}^{T-h-1} X_\tau\right) - c\right] \tag{51}$$

Step 3 from steps 1 and 2 we conclude by backward induction that Property 1 is true for any cohort t (with $1 \leq t \leq T-1$).

Appendix B: Proof of Lemma 2

Applying Thales' theorem to triangles ABC and ADF leads to:

$$CF = \frac{n_t}{1+n_t} AF. \tag{52}$$

Actually, CF is the optimal output produced by cohort t, that is X_t, while AF is the maximal quantities cohort t can produce to generate non-negative profit (equal to the difference between the perfect competition equilibrium supply and the output already produced by the previous cohorts). The property above can be rewritten as:

$$X_t = \frac{n_t}{1+n_t}(X_t + AC), \quad \text{or equivalently } AC = \frac{X_t}{n_t} = x_t. \tag{53}$$

Notice that AC is also the maximal quantities cohort $t+1$ can produce to generate non-negative profit. Then, equality (52) applied to cohorts t and $t+1$ becomes:

$$X_{t+1} = \frac{n_{t+1}}{1+n_{t+1}} AC, \quad \text{leading to } \frac{X_{t+1}}{n_{t+1}} = x_{t+1} = \frac{1}{1+n_{t+1}} x_t. \tag{54}$$

By backward induction, it turns out that:

$$x_t = \eta_{1,t} x_1, \quad \text{where } \eta_{1,t} \equiv \prod_{\tau=2}^{t} \frac{1}{1+n_\tau}. \tag{55}$$

References

Amir R, Grilo I (1999) Stackelberg versus Cournot equilibrium. Games Econ Behav 26:1–21

Anderson SP, Engers M (1992) Stackelberg versus Cournot equilibrium. Int J Ind Organ 10:127–135

Boyer M, Moreaux M (1986) Perfect competition as the limit of a hierarchical market game. Econ Lett 22:115–118

Carlton D, Perloff JM (1994) Modern industrial organization. HarperCollins, New York
Daughety A (1990) Beneficial concentration. Am Econ Rev 80:1231–1237
De Quinto J, Watt R (2003) Some simple graphical interpretations of the Herfindahl-Hirshman index. Mimeo
Etro F (2008) Stackelberg competition with endogenous entry. Econ J 118:1670–1697
Heywood JS, McGinty M (2008) Leading and merging: convex costs, Stackelberg, and the merger paradox. South Econ J 74:879–893
Julien LA, Musy O (2011) A generalized oligopoly model with conjectural variations. Metroeconomica 62:411–433
Lafay T (2010) A linear generalization of Stackelberg's model. Theor Decis 69:317–326
Matsumura T (1999) Quantity-setting oligopoly with endogenous sequencing. Int J Ind Organ 17:289–296
Pal D, Sarkar J (2001) A Stackelberg oligopoly with nonidentical firms. Bull Econ Res 53:127–135
Sherali HD (1984) A multiple leader Stackelberg model and analysis. Oper Res 32:390–404
Stackelberg H (1934) Marktform und Gleichgewicht. Springer, Berlin
Vives X (1999) Oligopoly pricing. Old ideas and new tools. MIT Press, Cambridge
Watt R (2002) A generalized oligopoly model. Metroeconomica 53:46–55